Universitat Politècnica de Catalunya
Facultad De Informática de Barcelona (FIB)

El uso de Blockchain para resguardar Analíticas de Aprendizaje.

MÁSTER INGENIERÍA INFORMÁTICA

TRABAJO DE FIN DE MÁSTER

Autor: Eduardo Rodríguez Ruiz

Director: Marc Alier Forment
Comité de tesis: Manel Frigola Bourlon
Karina Gilbert Oliveras
Jordi Delgado Pin

Fecha de la Defensa

24 de abril de 2019

Abstract

El uso de Blockchain para resguardar Analíticas de Aprendizaje.

por Eduardo Rodríguez Ruiz

Las analíticas de aprendizaje son una herramienta de gran utilidad ya que permiten mejorar y personalizar la manera en la que un estudiante llega a aprender. Sin embargo, se requiere del análisis de los datos personales del estudiante para lograr este fin. Esto se torna en una problemática delicada, ya que se requiere que la plataforma que desea hacer el análisis, provea al estudiante de mecanismos que aseguren el manejo cuidadoso de sus datos. Ademas, los datos deben de estar protegidos y se le debe permitir al estudiante controlar quién puede acceder a sus ellos. El propósito de este trabajo radica en analizar la viabilidad de usar la tecnología de Blockchain en conjunción a los mecanismos de seguridad informática implementados en la plataforma de analíticas de aprendizaje "Student Progress Snapshot", para darle más confianza a los alumnos por parte de la plataforma. Es por esto que el sistema basado en Blockchain, debe permitir a los usuarios tener la gestión de los permisos de acceso a su información personal de una forma clara, concisa y que además ayude a cumplir con las regulaciones dadas por la GDPR. Cabe señalar que la intención de este trabajo no es dar un sistema completo de seguridad basado en Blockchain, sino que solamente se abordará la gestión de accesos a la información personal.

Palabras Clave: Blockchain, Ethereum, Analíticas de Aprendizaje, Privacidad informática

Agradecimientos

Agradezco de manera inconmensurable a mis padres por permitirme seguir mi carrera académica, a mi hermano por ser inspiración para incursionar en una maestría y por los consejos dados. Sin ellos, esto no podría haber sido siquiera imaginado. De igual manera, agradezco a los tutores Marc Alier y Daniel Amo, ya que me ayudaron ampliar un poco más los horizontes, además de dar retroalimentación precisa durante todo el trabajo.

También, agradezco a mis amigos, el grupo de Smoxianos, los Perros Barceloneses y los integrantes de Mausoleo Barroco, por su apoyo, palabras de aliento y acompañarme en esta odisea, a Dani, por ofrecer su ayuda a pesar de no tener tiempo ni de respirar, a Yamile, por alentarme a terminar esta memoria y a Cinthya, por impulsarme a ser siempre una persona mejor cada día.

Finalmente, agradezco a mi perro por estar a mi lado durante la realización de este trabajo.

Índice general

Resumen I

Agradecimientos II

List of Figures V

1. Introducción 1
- 1.1. Introducción . . . 1
- 1.2. Motivación . . . 2
- 1.3. Hipótesis . . . 3
- 1.4. Objetivos . . . 4
  - 1.4.1. Objetivos Generales . . . 4
  - 1.4.2. Objetivos Particulares . . . 5
  - 1.4.3. Alcance . . . 5
- 1.5. Planificación . . . 6
- 1.6. Presupuesto . . . 7
- 1.7. Organización de la Tesis . . . 7

2. Marco Teórico. 9
- 2.1. Conceptos . . . 9
  - 2.1.1. Blockchain . . . 10
    - 2.1.1.1. Transacciones . . . 10
    - 2.1.1.2. Servidor . . . 12
    - 2.1.1.3. Proof of work . . . 12
    - 2.1.1.4. Red . . . 13
    - 2.1.1.5. Privacidad . . . 14
    - 2.1.1.6. Scripting . . . 14
  - 2.1.2. Ethereum . . . 15
    - 2.1.2.1. Cuentas de Ethereum . . . 16
    - 2.1.2.2. Transacciones de Ethereum . . . 16
    - 2.1.2.3. Interacción con Ethereum . . . 17
  - 2.1.3. Smart Contracts . . . 18
    - 2.1.3.1. Usos de Smart Contracts . . . 18
  - 2.1.4. Learning Analytics . . . 19

2.1.4.1. Moodle 20
2.2. Privacidad y Protección de Datos 21
2.2.1. GDPR 22
2.2.1.1. Principios 22
2.2.1.2. Derechos del interesado 23
2.2.1.3. Responsable del Tratamiento y Encargado del Tratamiento 25
2.2.1.4. Delegado de Protección de Datos 26
2.2.2. Privacidad y Analíticas de Aprendizaje 27
2.3. Trabajos Relacionados 29
2.3.1. BBLAP 29
2.3.2. Hawk 31
2.3.2.1. Seguridad 31
2.3.3. Desventajas de los trabajos relacionados 33
2.4. Student Progress Snapshot 34

3. Trabajo Realizado **36**
3.1. Infraestructura 36
3.2. Permisos usando Smart Contracts 37
3.3. Estructura 38
3.4. Implementación 40
3.4.1. Implementación de Etherum 40
3.4.1.1. Cuentas de Ethereum 42
3.4.1.2. Smart Contract Desarrollado 43
3.4.2. Middleware 45

4. Conclusiones **47**
4.1. Conclusiones 47
4.2. Trabajo a futuro 48

Bibliografía **49**

Índice de figuras

1.1. Diagrama de Gantt con tiempos de implementacion 6

2.1. Transacciones en la red. Fuente: [1] 11
2.2. Creación de Hashes dentro de la red [1]. 12
2.3. Contenido de los Bloques. Fuente: [1] 13
2.4. Modelo de privacidad antiguo vs Modelo de privacidad de Blockchain. Fuente: [1] 14
2.5. Diseño de la plataforma BBLAP. Fuente: [2] et al 2018 30
2.6. Contrato de Hawk. Fuente: Kosba et al 2016 32
2.7. Esquema de la plataforma Student Progress Snapshot. 35

3.1. Esquema de la interacción entre el Blockchain y el Student Progress Snapshot. 37
3.2. Ejemplo de contenido de un bloque Génesis. Fuente: [3] 41
3.3. Comando para ejecución de la red Blockchain. 41
3.4. Nodo de Ethereum ejecutándose. 43

Capítulo 1

Introducción

1.1. Introducción

En la actualidad, se vive en la era de la información, donde se genera a cada segundo millones de datos, no solo por académicos ni los diferentes tipos de industrias, sino principalmente por la población mundial. Toda esta generación de información se debe gracias a la multifuncionalidad que ofrecen los sistemas informáticos desde las redes sociales, la captura de fotos, huellas digitales, datos biométricos, hasta cuentas bancarias, entre muchos datos más. Al existir tanta información importante en sistemas informáticos, existe un valor abstracto en dicha información, lo cual hace que la información se vuelva un recurso valioso y por lo tanto, requiera de ser protegido.

La seguridad informática, es el área encargada de proteger los datos computacionales, de tal manera que se cumplan diversos objetivos, entre los que se encuentran la confidencialidad, que implica quién tiene acceso a la información; e integridad, que significa que la información no debe ser corrompida y solo puede ser editada por los usuarios autorizados [4]. Para lograr estos objetivos, existen algoritmos de encriptación que dificultan o imposibilitan la visualización de los contenidos de los datos si no se poseen las llaves indicadas. Ademas, existen soluciones de autenticación, que validan que una entidad sea

quien dice ser; autorización, que revisan los permisos de los usuarios autenticados; y contabilidad, que guardan información acerca de que acciones se llevaron a cabo en el sistema por los usuarios [4].

Hasta hace poco, las compañías y plataformas que capturaban y preservaban la información de las personas físicas, solamente tenían como enfoque de seguridad los accesos no permitidos a la información que guardaban. Sin embargo, se omitían los derechos del usuario cuya información estaba siendo guardada, los cuales no tenían jurisdicción sobre sus datos. Con el Reglamento General de la Protección de Datos (GDPR por sus siglas en inglés) [5], se refuerzan los derechos de los usuarios, dándole los permisos de autorización a ellos y no a los que mantienen la información. Esta situación es delicada, ya que hasta el momento las plataformas responsables de guardar la información de los usuarios no poseen mecanismos que le brinden al usuario la posibilidad de dar permisos respecto quién puede ver su propia información, ni de asegurar que dichos permisos se mantendrán íntegros.

Las Análiticas de Aprendizaje (Learning Analítics, en inglés) son una tecnología que provee a los docentes una forma de mejorar su método de enseñanza, mediante la captación de información de cada alumno. Dicha captación permite que la enseñanza sea personalizada para que el alumno obtenga el mayor beneficio de ella. Debido a la condición inherente del acceso de datos para su análisis, cualquier actor, ya sean instituciones educativas, gubernamentales o privadas que deseen implementar analíticas de aprendizaje, tienen que manejar la información que poseen de la manera más cautelosa, preservando los derechos de los alumnos respecto a su información.

1.2. Motivación

Las analíticas de aprendizaje son un paradigma revolucionario tanto para los estudiantes como para los profesores [6]. Sin embargo, éstas se enfocan en el manejo de los datos de

los estudiantes, que en muchos casos llegan a ser menores de edad. Dado este motivo, existen las preocupaciones del impacto que tiene el análisis de los usuarios respecto a su privacidad. Un claro ejemplo de éstas preocupaciones existen en el caso de la plataforma de análiticas de aprendizaje InBloom, la cual, debido la cantidad de datos personales de los alumnos que se tenía en la plataforma, los padres objetaron que no deseaban que su información estuviera a la venta a terceros. Debido a este temor, la plataforma cerró sus operaciones por instrucción del gobierno de los Estados Unidos de América [7, 8]. Actualmente, no existe ningun tipo de solución tecnológica para este problema. Generalmente, las soluciones para plataformas de analíticas de aprendizaje estan centradas en la interoperabilidad de datos dentro de las plataformas, como las soluciones IMS Caliper, o X-API. Sin embargo, estas soluciones no abordan las problemáticas hacia las analíticas de aprendizaje arraigadas en factores humanos, tales como la desconfianza, la angustia, el escepticismo, malentendidos, entre otros.

1.3. Hipótesis

Los objetivos de la seguridad informática, se enfocan en proponer y emplear contramedidas hacia amenazas externas, por ejemplo, hackers, piratas cibernéticos, entre otros, asumiendo que la información mantenida internamente está segura e ignorando el factor humano. El enfoque utilizado para el desarrollo de este trabajo radica en que las plataformas tecnológicas que captan y guardan información de usuarios, en específico las de análiticas de aprendizaje, fungen como actores internos y externos respecto a los datos de sus usuarios. Estas plataformas deben de garantizar que la información que ellos almacenan solamente puede ser accedida por agentes que el usuario dió el permiso explícitamente, de tal forma de que si el usuario no le permite el acceso a su información a miembros internos de la plataforma o a toda la plataforma en sí, la plataforma solo debe de almacenar la información.

El Blockchain es la tecnología en la cual se pueden hacer transacciones electrónicas sin la necesidad de que exista confianza entre los individuos [1]. Si asumimos que el usuario carece de la confianza hacia la plataforma y no se desea que se incluya un intermediario, una implementación basada en Blockchain, en la que interactúen usuarios, los encargados de la plataforma responsables de la preservación de los datos de los usuarios y miembros que requieran del uso de la información de los usuarios para la mejora del aprendizaje del usuario y terceros, proveerá la confianza hacia la plataforma mientras las transacciones sean los permisos de los datos del usuario.

Ésta hipótesis será aplicada al proyecto de análiticas de aprendizaje "Student Progress Snapshot" del doctorante Daniel Amo y del Dr. Marc Alier, el cual permitirá a los profesores analizar las actividades de sus estudiantes dentro de los cursos de la plataforma Moodle.

1.4. Objetivos

1.4.1. Objetivos Generales

En este proyecto se buscan cumplir con los siguientes objetivos:

- Crear un mecanismo tecnológico que permita guardar los permisos de autorización a la información guardada y, que los permisos no sean fácilmente manipulados.
- Otorgar un mecanismo claro para la gestión de la autorización de información a los usuarios de la plataforma de análiticas de aprendizaje.
- Determinar si el uso de la tecnología Blockchain resuelve el problema de falta de confianza hacia las plataformas de analíticas de aprendizaje.

1.4.2. Objetivos Particulares

1. Tener una solución de Blockchain sin la necesidad de un coste económico a la infraestructura ya creada.

2. Crear smart contracts encargados de mantener los permisos de los docentes respecto a la información de los usuarios y que dichos permisos tengan expiración.

3. Implementar middleware capaz de vincular tanto la plataforma de analíticas de aprendizaje con los smart contracts dentro del Blockchain, de tal manera que la plataforma pueda añadir y obtener los permisos de los docentes.

1.4.3. Alcance

Este trabajo entra dentro del contexto de aplicar a la plataforma de analítica de aprendizaje "Student Progress Snapshot" diversos controles de seguridad, con el fin de proteger la información personal de los alumnos. Dado el alcance de los objetivos generales, este trabajo se centra en dotar a dicha plataforma de mecanismos para la autorización y la contabilidad acerca de quién puede acceder y modificar dicha información, con ciertas restricciones evidenciadas en los objetivos particulares. Queda fuera del alcance de este trabajo proveer una solución de seguridad completa, ya que los mecanismos de autenticación para verificar las identidades de las personas que intenta obtener autorización para acceder a la información; así como los mecanismos de cifrado que se pueden usar para preservar la confidencialidad ante atacantes externos, no son objeto de estudio de este trabajo. Sin embargo, al cumplir los objetivos generales y específicos, se está haciendo una aportación a la seguridad del sistema en general.

1.5. Planificación

La planificación de este proyecto se dividió en 4 fases: Aprendizaje, investigación, diseño y finalmente, desarrollo. La fase de aprendizaje consistió en aprender sobre cómo funcionaba la tecnología Blockchain y los usos más comunes que tenía, además de en qué consistían los smart contracts, cómo interactúan con la Blockchain y qué limitaciones encontraban, además de una ligera introducción referente al tema de las analíticas de aprendizaje . En la fase de investigación, se enfocó a indagar sobre las implicaciones que conlleva no tener un sistema seguro tanto en los ámbitos legales como sociales, profundizando en las responsabilidades que se tienen al tener que manejar con información; también se buscaron trabajos relacionados que hayan usado la tecnología Blockchain para proponer una solución al problema del manejo de datos personales y determinar si la solución cumple con todas las problemáticas o en su defecto, identificar las áreas de oportunidad de dichos trabajos. Durante la fase de diseño, se tomó como base la investigación previa y los objetivos del trabajo para proponer una solución en la cual el smart contract contenga la información necesaria para determinar a qué información el docente y la plataforma tienen derecho de acceder y que el alumno tenga mayor control respecto a su privacidad. Finalmente, la fase de desarrollo se dedicó a implementar el diseño propuesto y en caso de que se haya ignorado u omitido un aspecto importante para la gestión de acceso a la información del alumno, corregir el diseño y que ese cambio se vea reflejado en la implementación. El cumplimiento de dichas fases se hizo a través de 4 meses, consistiendo en dos semanas para la fase de aprendizaje, 5 semanas para la fase de investigación, dos semanas para la fase de diseño y 7 semanas para la fase desarrollo. Adicionalmente, se dispuso de 2 semanas para el desarrollo del escrito de este trabajo.

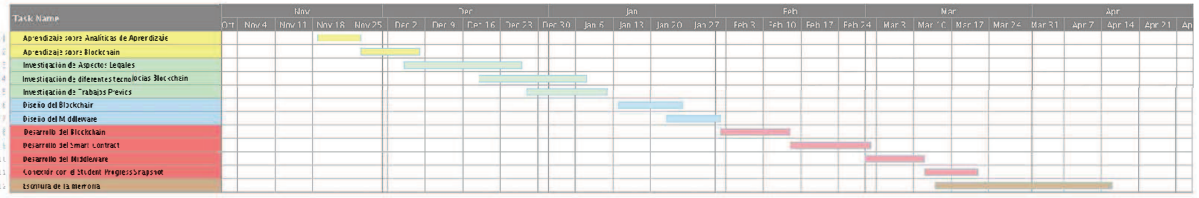

FIGURA 1.1: Diagrama de Gantt con tiempos de implementacion

1.6. Presupuesto

Para el presupuesto destinado para de este proyecto, como aspectos a tomar en cuenta se encuentran el equipo de desarrollo, tanto de personal como tecnológico. A continuación se hace un desglose de los gastos requeridos para que se lleve a cabo el proyecto. Cabe señalar que el único gasto fijo que existe es el del computador para el desarrollo, siendo este equipo un MSI GE62 Apache pro de un valor aproximado de 1100 euros, mientras que el gasto recurrente sería el salario del desarrollador, que se estipula en 1000 euros, siendo este el sueldo pasado del desarrollador. Siendo el total de 5100 euros destinado al mero desarrollo del sistema.

Este presupuesto es ajeno al mantenimiento mensual del sistema ya en producción, debido a que se debe tener en cuenta tanto el coste que implica tener tanto al servidor middleware, como el nodo inicial que mantenga la Blockchain ejecutándose de manera constante, dada la necesidad de preservar la disponibilidad del sistema. Con esto en consideración, se debería tener un equipo computacional capaz de tener tanto el middleware como el nodo de la red Blockchain en sí .Para este equipo, se hizo una cotización en el sitio web de Dell y al analizar las opciones posibles, el Dell EMC PowerEdge T440 cumple con los requerimientos ofrecidos ya que este ofrece hasta dos procesadores Intel Xeon, 1 Terabyte de memoria RAM, una tarjeta gráfica NVIDIA Quadro P4000 de 8 Gigabytes de memoria, 4 Discos duros de 96 Terabytes con un precio de 4050 euros.

1.7. Organización de la Tesis

Para el desarrollo de esta tesis se comienza introduciendo al lector los conceptos necesarios de Blockchain,smart contracts, análiticas de aprendizaje y privacidad, así como el contexto que existe respecto a las soluciones que se han propuesto en el área de privacidad y Blockchain. Posteriormente,se describe el proyecto “Student Progress Snapshot” (SPS) y se desarrolla a fondo la propuesta descrita, ahondando en las tecnologías usadas, además

de proporcionar las ventajas que tienen. Finalmente, se darán las conclusiones y una propuesta de trabajo a futuro.

Capítulo 2

Marco Teórico.

Para comprender a fondo la relevancia del trabajo aquí desarrollado, es necesario explicar los conceptos que abarca la privacidad basada en Blockchain, además de los trabajos realizados que proponen una solución similar, sus aportaciones y las áreas de oportunidad que tiene. Este capítulo se enfocara en dar una descripción de la infraestructura de una red Blockchain, el sistema Student Progress Snapshot, las analíticas de aprendizaje, las regulaciones gubernamentales respecto al manejo de datos por particulares. Posteriormente, se observará el estado del arte, las aportaciones que proveen las soluciones actuales y las deficiencias al abordar el problema.

2.1. Conceptos

Para comprender como proveer un mecanismo de confianza para la asignación de permisos respecto los datos de los usuarios mediante el Blockchain, es necesario comprender ciertos conceptos básicos, tales como la infraestructura del Blockchain, qué es un smart contract, cómo opera y cómo funcionan las analíticas de aprendizaje. En esta sección se explican estos conceptos y la relación entre ellos.

2.1.1. Blockchain

Originalmente creado por Satoshi Nakamoto en 2008 como respuesta al creciente comercio por internet y, de cómo este comercio dependía casi en su totalidad de entidades financieras fungiendo de terceros de confianza encargados de procesar las transacciones electrónicas; Nakamoto propone un sistema en el cual dos individuos puedan hacer transacciones entre ellos sin la necesidad de ningún intermediario de confianza. Este sistema se basa en pruebas criptográficas, las cuales tienen la propiedad de ser difíciles de poder ser revertidas, protegiendo de esta manera a los vendedores; de igual forma, para proteger a los compradores se podrían implementar rutinas de fideicomisos. Además, para resolver el problema del double spending, el cual implica que no se puede usar la misma moneda para hacer dos pagos distintos, se utiliza un servidor distribuido peer-to-peer y, para que genere las transacciones por un orden cronológico, también posea una marca de tiempo (timestamp, en inglés). Este sistema ofrece que sus transacciones permanezcan seguras si la mayoría de los nodos encargados del procesamiento se mantienen honestos, o en otras palabras, mantienen las transacciones pasadas de manera correcta y no poseen transacciones que nunca fueron hechas. A este sistema Nakamoto lo nombró Bitcoin, por la moneda electrónica que se usa para las transacciones, además, al método en el que funciona la red distribuida y la forma en la que se guardan las transacciones, se le conoce como Blockchain.

2.1.1.1. Transacciones

El principio de una red Blockchain yace en las transacciones que, como se mencionó anteriormente,usa una moneda electrónica como forma de intercambio; se le denomina el término de moneda electrónica a una cadena de firmas digitales. Para poder hacer una transacción de una de éstas monedas electrónicas, el dueño actual de la moneda tiene que firmar digitalmente esa moneda mediante un hash de la transacción anterior en la que esa moneda fue usada y la llave pública del nuevo dueño al final de la cadena de la moneda.

De esta forma, el cobrador puede verificar las firmas digitales de la cadena para saber el historial de dueños.

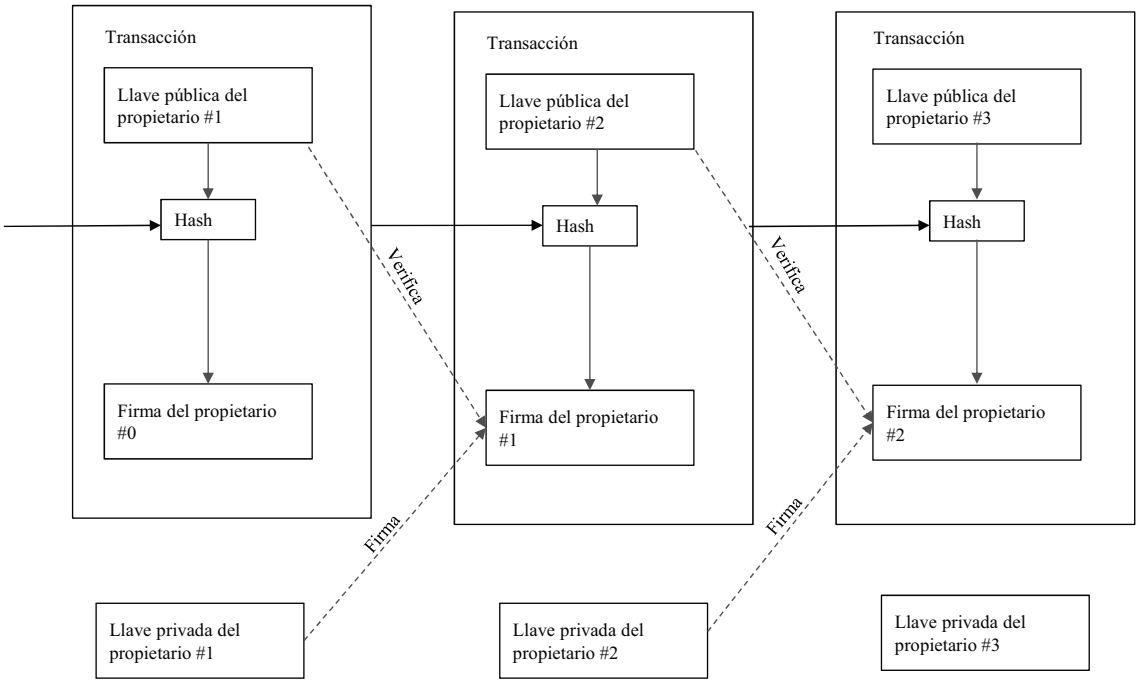

FIGURA 2.1: Transacciones en la red. Fuente: [1]

El problema de esta implementación es que el cobrador no tiene la certeza de que un dueño anterior no haya transaccionado dos veces con la misma moneda. La solución que se usa en los métodos tradicionales de transacciones monetarias implica una autoridad central de confianza que se dedique a acuñar las monedas y, cuando se hace una transacción, esa moneda regresa a la autoridad central para que se vuelva a acuñar, de esta manera, una moneda no puede ser usada para dos transacciones distintas. Es por esto que para el sistema no tenga ningún tercero de confianza y que el cobrador sepa que los dueños pasados no hayan usado esa moneda para dos transacciones, solo se considera la transacción más temprana de esa moneda, descartando las transacciones posteriores del dueño que esa moneda cambie de dueño. Es por este motivo que todas las transacciones sean visibles para todos los nodos y, que los nodos de la red estén de acuerdo en el historial único en el que las transacciones fueron hechas.

A manera para incentivar a los nodos a mantenerse en la red, en la primera transacción de un nuevo bloque se crea una moneda y se le entrega al creador del bloque. Esto, provee un método de introducir monedas a circulación, ya que no se posee ninguna autoridad central dedicada a emitirlas. La adición de monedas hace símil a la situación de los mineros de oro al insertar dicho metal a la circulación, con la diferencia que los recursos que se utilizan son el poder de cómputo y la energía.

2.1.1.2. Servidor

Un elemento vital de esta solución recae en el servidor timestamp. Este funciona tomando un hash de un bloque de objetos a los cuales se les debe marcar el tiempo y anunciar públicamente el hash, de esta manera, se demuestra que el objeto existió durante ese tiempo para que se pudiera obtener su hash. Cada marca temporal incluye la marca temporal anterior en su hash, de esta forma, se crea una cadena de bloques, con cada marca temporal reforzando las marcas temporales pasadas.

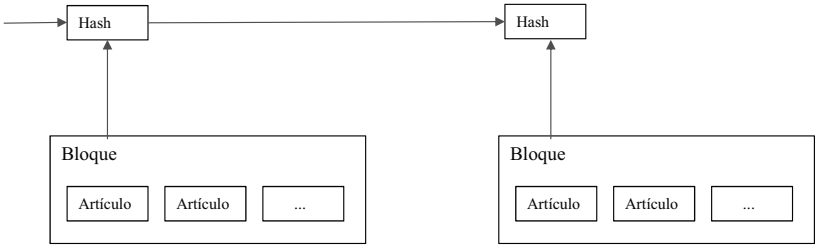

FIGURA 2.2: Creación de Hashes dentro de la red [1].

2.1.1.3. Proof of work

Para la implementación de un servidor timestamp peer-to-peer distribuido, se requiere de un proof-or work, el cual implica buscar un valor que, cuando se pasa a un hash, como con SHA-256, el hash resultante comienza con un número determinado de cero bits. El trabajo requerido en promedio es exponencial respecto al número de cero bits consecutivos.

La implementación del proof-of-work para este servidor comienza incrementando el valor nonce del bloque, el cual solo puede ser usado una vez, hasta que se encuentre que el hash del bloque tenga los cero bits consecutivos requeridos. Una vez que el esfuerzo de procesamiento,también llamado minado, se haya hecho para cumplir el proof-of-work, ese bloque no puede ser cambiado sin que se tenga que hacer ese trabajo computacional y, en caso de que existan bloques encadenados después de este, se tendría que hacer también el trabajo computacional de los bloques posteriores. Es por este motivo, que si la mayoría de los nodos controlan el poder de procesamiento del servidor distribudo, la cadena crecerá de manera más acelerada y, dado a que siempre se escoge la cadena más grande como la cadena principal, el resto de las cadenas no serán aceptadas; además, si un atacante deseara modificar un bloque, debe primero de modificar los bloques posteriores al que tiene fijado, además de alcanzar la cadena principal y superar el poder de computo de todos los nodos honestos, lo cual requiere que el atacante poseea un poder de procesamiento extremadamente potente para poder hacer todo el proof-of-work necesario para apoderarse de la cadena.

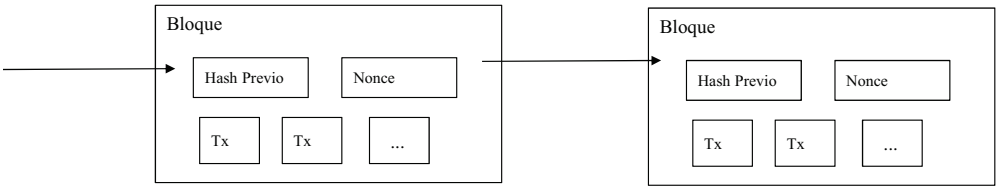

FIGURA 2.3: Contenido de los Bloques. Fuente: [1]

2.1.1.4. Red

La red funciona de la siguiente manera:

1. Las transacciones nuevas se envían a todos los nodos.
2. Cada nodo agrega las transacciones al bloque.
3. Cada nodo trabaja en encontrar un proof-of-work para su bloque. A estos nodos se les conoce coloquialmente como mineros.

4. Cuando el nodo encuentra el proof-of-work válido, envía el bloque a todos los nodos.

5. Los demás nodos verifican que el bloque sea correcto si las transacciones son válidas.

6. Si es correcto el nodo, se añade el bloque nuevo a su cadena, usando el hash del bloque aceptado como el hash pasado.

2.1.1.5. Privacidad

La manera tradicional en la que se hacen las transacciones fuera de Blockchain, limita el acceso de la información a los involucrados y a un tercero de confianza. Sin embargo, por definición, se excluye este método dada la necesidad de anunciar cada transacción de manera pública. Afortunadamente, la privacidad se logra manteniendo las llaves públicas anónimas, de tal forma, se puede saber que se está creando una transacción entre dos individuos y, el contenido de la transacción, pero sin saber qué individuos están involucrados.

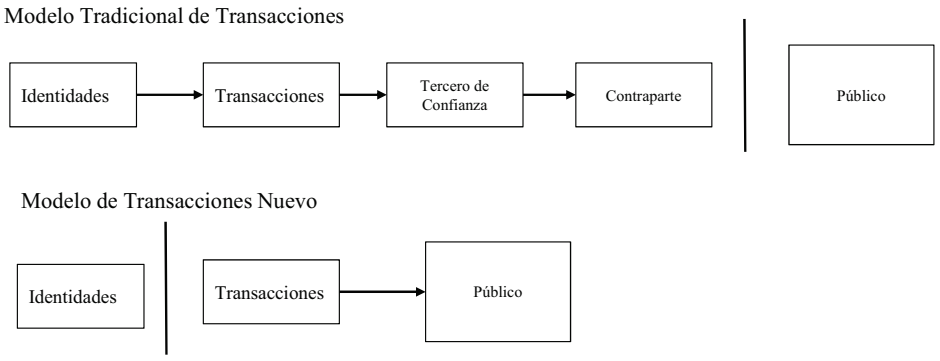

FIGURA 2.4: Modelo de privacidad antiguo vs Modelo de privacidad de Blockchain. Fuente: [1]

2.1.1.6. Scripting

Un punto que se debe de recalcar es, que cada transacción para la transferencia de bitcoins entre individuos está ligada a un pequeño script escrito en el lenguaje para Bitcoin. Estos scripts fungen como mediador entre las entradas y salidas de las transacciones,

por ejemplo, la función para liberar fondos, recibiendo como entrada la llave pública asociada a una dirección de bitcoin y como resultado da el permiso de liberación. Para evitar complejidades computacionales como lo son los bucles infinitos o el problema de paro, el lenguaje de scripting de bitcoin no es Turing-completo, por lo tanto, es menos expresivo [9].

Acorde a Buterin *et al.* [10] esta simpleza del lenguaje de scripting hace que tenga ciertas limitaciones, como por ejemplo:

- Dada su necesidad de no tener bucles infinitos, no existe la posibilidad de hacer bucles, por lo tanto, el lenguaje no es Turing-completo
- No existe manera de que el script provea una manera de retirar dinero de manera precisa. Esta limitación se llama Value Blindness.
- Dado a que los bitcoins solo tienen los estados de gastados y no gastados, no se pueden crear scripts capaces de tener contratos multicapas, por lo que solo se pueden crear contratos sencillos. Buterin *et al.*nombran esta falta como Lack of State
- Las monedas desconocen los datos que constituyen a la Blockchain, tal como el hash previo o el valor nonce. A esto se le denomina Blockchain-Blindness.

2.1.2. Ethereum

Descrito en el Ethereum's White paper de Buterin *et al.* [10]: "*la intención de ethereum es combinar y mejorar a partir de los conceptos de scripting, monedas virtuales alternativas y permitir a los desarrolladores crear aplicaciones basadas en consensos arbitrarios que tienen la escalabilidad, estandarización, feature-completeness, facilidad de desarrollo e interoperabilidad que ofrecen estos paradigmas al mismo tiempo. Etherum logra esto al construir lo que es esencialmente la capa abstracta idónea: una Blockchain con un lenguaje de programación Turing-complete incorporado, habilitando la capacidad para que*

cualquiera escriba smart contracts y funciones de estado transaccionales. [...] Los Smart Contracts, que son 'cajas' criptográficas que se abren solamente si ciertas condiciones se cumplen, también se pueden construir encima de la plataforma, con más poder que es ofrecido por el scripting de Bitcoin gracias a las capacidades conjuntas de saber del estado del blockchain,value awareness, ser Turing-completo y poseer estados."

2.1.2.1. Cuentas de Ethereum

Se define como cuentas a los objetos que constituyen un estado, cada cuenta teniendo una dirección de 20 bytes y transiciones de estado, éstas siendo transferencias directas de valores e información entre cuentas. Cada cuenta posee los siguientes 4 valores:

- Un nonce, usada para asegurar que una transacción solo se puede hacer una vez.
- El balance de la moneda de la cuenta, ether.
- Si existe, el código de contrato.
- El almacenamiento interno de la cuenta.

En ethereum, se clasifican las cuentas en dos tipos: las de propietarios externos y las cuentas contractuales. La diferencia yace que en que la primera es manejada mediante llaves privada, mientras que las cuentas contractuales son controladas mediante su propio código y, cada vez que recibe un mensaje, su código se activa, permitiendo que se lea y se escriba en su almacenamiento interno.

2.1.2.2. Transacciones de Ethereum

Este término se refiere a paquetes cifrados de datos que guardan algún mensaje enviado desde una cuenta propietaria externa. Éstas contienen: el receptor del mensaje, la firma del emisor, la cantidad de monedas, llamadas ether, que se envía, un campo de datos opcional,

el valor "Start gas" que representa el número máximo de cálculos que la ejecución de la transacción puede hacer y, un valor "Gas Price" representando la cuota el emisor debe de pagar por cálculo computacional. Los valores de receptor, firma y la cantidad monetaria son comunes en cualquier implementación de criptomonedas. Por otra parte, los valores de *Start Gas* y *Gas Price* son vitales para Ethereum, ya que con ellos se prevé que no existan bucles infinitos accidentales u hostiles; cada transacción requiere tener un límite de cuántos cálculos puede hacer.

La unidad fundamental computacional es el "gas", y, usualmente los cálculos computacionales requieren de 1 gas para ejecutarse, pero las operaciones que requieran de más poder de cálculo, llegan a costar 5. Además, se paga una cuota de 5 gas por cada byte de la transacción. Este método existe por motivos de seguridad a la red, ya que el atacante tendría que pagar por cada recurso que consuma, tal como el ancho de banda, el almacenamiento de datos y el poder computacional.

Existe un elemento similar a las transacciones, llamada mensaje, que se comporta de manera similar, que contiene el receptor y emisor del mensaje, la firma del emisor, la cantidad de ether que se quiere enviar, un campo opcional y un valor de *Start Gas*. La diferencia entre un mensaje y una transacción es que el mensaje es producido por un contrato, de esta manera, un contrato se puede comunicar con otros contratos.

2.1.2.3. Interacción con Ethereum

Para poder interactuar con la Blockchain de Ethereum y con sus contratos, los nodos de Ethereum ofrecen una interfaz en la cual el nodo puede llamar a los procedimientos remotos como si fueran ejecutados de forma local, denominada un RPC (Remote Procedure Call, por sus siglas en inglés) [11] y está disponible a través de HTTP e IPC. Dado que usar la interfaz RPC resulta en un proceso tedioso y propenso a los errores, se desarrolló Web3.js, una librería que funciona sobre la RPC que provee una interfaz más amigable para su uso y sin ser tan propensa a los errores [12].

2.1.3. Smart Contracts

Los smart contracts son scripts de código definidos por los usuarios que especifican reglas que rigen a las transacciones dentro de una red de Blockchain y, que se ejecutan dentro de la misma. Cuando se manda a llamar un contrato, su código es ejecutado por los nodos de la Blockchain, que llegan al consenso descentralizado del resultado de la ejecución y actualizan la Blockchain; este proceso ocurre cada vez que el contrato recibe una llamada, ya sea de una cuenta de propietario externo u otro contrato. Un aspecto vital que se debe de considerar es que el contrato se considera como un tercero de confianza, pero que este tercero provee la confianza para la disponibilidad e integridad de la transacción, más no de su privacidad, ya que el estado del contrato es visible a todo el público [13]. Esto implica que no se recomienda que un smart contract contenga información personal, ya que todos los nodos reciben esa información al momento de minar la transacción y añadirla a un bloque.

2.1.3.1. Usos de Smart Contracts

En [14] se presenta un análisis de los usos más comunes de los smart contracts a través de las Blockchains públicas, los cuales son:

Financieros El uso más común de los smart contracts, éstos se dedican a administrar, recolectar o distribuir dinero. Algunos de estos contratos certifican la posesión de algún bien en el mundo real, endosando su valor y manteniendo el historial de sus transacciones.

Notariales Dada la naturaleza de inmutabilidad de la Blockchain, estos contratos guardan datos de manera permanente, para así certificar su autoría y procedencia. Algunos de estos contratos permiten a los usuarios guardar el hash de un documento en el Blockchain, de tal manera demostrar su la existencia del documento.

Juego Aquí se agrupan contratos que implementan juegos de apuesta, ya sean de azar, como los dados, la lotería, ruleta, entre otros más; juegos de habilidad y una mezcla de ambos.

Billetera electrónica Este tipo de contratos fungen como intermediarios que simplifican la interacción con la Blockchain. se encargan de guardar llaves, manejar dinero y otros contratos.

Librería Estos son contratos que poseen operaciones de uso general (como operaciones matemáticas o de manipulación de strings), usadas principalmente por otros contratos

2.1.4. Learning Analytics

La primera conferencia en ánaliticas de aprendizaje y conocimiento (LAK 2011, por sus siglas en inglés) define a las analíticas de aprendizaje como la medición, recolección, análisis y reportes de datos de los estudiantes y sus contextos, con los propósitos de entendimiento y la optimización del aprendizaje y el ambiente en el que ocurre. Por otra parte, ésta definición abarca la mayoría del área de la investigación didáctica, es por esto que las análiticas de aprendizaje tienen los siguientes supuestos: las analíticas de aprendizaje hacen uso de datos pre-existentes y legibles por máquinas, además de que sus técnicas pueden ser usadas para gestionar big data, cúmulos intensos de datos que resultaría impráctico manejarlos manualmente.

Acorde a [15] existen tres motores que motivan el surgimiento y desarrollo de las análiticas de aprendizaje y sus campos relacionados:

Big data la sociedad se enfrenta a un reto dado por la big data. Las empresas usan analíticas para extraer valor de los datasets inmensos,usandolos de para identificar patrones de comportamiento y desarrollar campañas de marketing. El amplio uso de ambientes de aprendizaje virtuales, como son Blackboard o Moodle implica que

las instituciones educativas también tienen que lidiar con grandes cúmulos de datos y, con el paso de los días, éstas plataformas añaden cantidades de datos personales, interactivos e información académica. A pesar de la información que poseen éstas plataformas, la explotación de esta información al analizarlas, resultan básicas

Aprendizaje en línea El aprendizaje en línea ofrece una gran línea de beneficios, sin embargo, igual conlleva problemas en comparación con el aprendizaje tradicional, por ejemplo, los estudiantes llegan a tener un sentimiento de aislamiento dada la falta de contacto con sus profesores y compañeros, además de desorientación en el espacio en línea, tener problemas técnicos o perder su motivación. De la misma manera, los docentes carecen de las pautas que señalan cuando los estudiantes se encuentran confundidos, ausentes, entre otros, además de tener problemas al interpretar y evaluar el aprendizaje y calidad de participación de los alumnos.

Preocupaciones políticas Existe una exigencia para que las instituciones educativas lleguen a medir, demostrar y mejorar su desempeño.

2.1.4.1. Moodle

En [16] se describe al sistema para la administración de cursos educativos Moodle como un software open-source orientado a ayudar a educadores comunidades educativas en línea. Esta plataforma ha sido instalado e implementada en múltiples organizaciones, instituciones educativas y universidades, y, dado a que Moodle provee el acceso completo a su código fuente, la organización puede hacer los cambios que ésta vea necesarios, por ejemplo: crear nuevos cursos o añadir contenido que involucre activamente a los estudiantes.

Todo esto se debe a que Moodle fue diseñado para apoyar el estilo de aprendizaje llamado pedagogía construccionista social. Este estilo de aprendizaje se basa en la creencia de que los estudiantes aprenden de mejor forma cuando interactúan directamente con el material de aprendizaje, lo cual implica que ellos puedan construir nuevo contenido didáctico y

que se puedan comunicar respecto el contenido con otros estudiantes, y si bien este es el estilo de aprendizaje que Moodle mejor se acopla, no es forzoso seguirlo.

Para poder personalizar la plataforma acorde a las necesidades de la organización, Moodle provee de múltiples herramientas, tanto estáticas, como interactivas y actividades donde los estudiantes puedan interactuar entre sí mismos.

Por otra parte, Moodle no solamente ofrece servicios para los alumnos, sino que también recauda la información detallada que un estudiante realiza y la guarda en una bitácora. Esta bitácora guarda cada click que el estudiante hace al navegar a través de la plataforma. Está información guardada puede ser filtrada por cursos,participantes, dias y actividades, de tal manera de que el instructor o docente pueda determinar qué estudiante ha participado en qué curso, qué hizo en el curso y cuándo lo hizo. Respecto a las actividades, tal como los exámenes, en la bitácora no solamente se guarda la calificación que obtuvo el estudiante, sino también el tiempo en el que se tomó al hacer ese examen y un análisis detallado de las respuestas. De igual manera, los instructores pueden obtener de manera sencilla los reportes generados, por ejemplo el reporte de las actividades individuales de un estudiante o los estudiantes inscritos en una actividad específica.

Todo esto es posible gracias a la implementación técnica de Moodle, la cual no guarda las bitácoras como archivos de texto, sino en una base de datos relacional, tal como lo es MysQL, PostgreSQL, Oracle, Access, entre otras. Es importante recalcar que el análisis de la información, no es hecha por los docentes involucrados en la plataforma, sino por los administradores que tengan conocimientos de Big Data y minería de datos.

2.2. Privacidad y Protección de Datos

Un aspecto vital de este trabajo está orientado a la protección y privacidad de datos, es por esto que se requiere hablar respecto a las legislaciones que existen para la preservación de

estos derechos y de qué pautas una plataforma de analíticas de aprendizaje debe cumplir para que pueda asegurar los datos que maneja.

2.2.1. GDPR

El Reglamento General de la Protección de Datos (GDPR por sus siglas en inglés) [5] está orientado a la protección de los derechos fundamentales de las personas físicas, en particular, el derecho a la protección de sus datos personales. También establece las reglas para la protección de las personas con respecto al proceso de sus datos personales, además de dar pautas para el libre movimiento de información personal. Esta regulación se aplica a cualquier procesamiento de información mientras el propietario o el encargado del tratamiento resida en la Unión Europea, independientemente de dónde se procesa la información.

2.2.1.1. Principios

Dentro de la GDPR, se presentan los principios que se deben acatar y la relación que tiene la persona que generó esos datos, llamado "interesado" (data subject, en inglés) y la entidad que maneja los datos, denominado "responsable del tratamiento" (controller, en inglés). A continuación, se mostrarán los principios más relevantes para este trabajo:

- Acorde al artículo 5, los datos personales deben ser recolectados para un propósito específico, explícito y legítimo. Dicha recolección es limitada en relación a la necesidad. Además, se debe de cerciorar que los datos son correctos y en caso de ser erróneos, tienen que ser rectificados o eliminados de manera inmediata. También, los datos se deben de mantener almacenados de tal manera que sean guardados hasta que ya no sean necesarios. Existen extensiones si los datos se procesan por motivos de archivación, interés público, académico o científico. Sobre todo, el encargado de tratamiento debe de asegurar que los datos personales se deben de procesar

de tal manera de que se garantice su seguridad, esto incluye la protección contra procesamiento no autorizado o ilegal y contra la pérdida de datos, para preservar su integridad y confidencialidad.

- El artículo 6 establece que el procesamiento de los datos personales es legal si el interesado da su consentimiento para un propósito específico, o si el procesamiento es necesario para los intereses particulares del responsable de tratamiento o un externo, exceptuando cuando esos intereses estén sobre los intereses o derechos fundamentales del interesado, particularmente si el interesado es un infante

- Respecto al consentimiento, el artículo 7 plantea que el responsable de tratamiento debe de poder demostrar de manera explícita el consentimiento del interesado, además de que ese consentimiento debe estar en un lenguaje legible. También, se define que el interesado es libre de remover su consentimiento cuando lo desee y debe de ser igual de sencillo poder quitar su consentimiento como lo es darlo.

- El artículo 8 explica que, cuando el interesado es un menor de edad, el procesamiento de sus datos solo será legal si su tutor legal consiente al procesamiento.

2.2.1.2. Derechos del interesado

En el capítulo 3 se definen los derechos que tiene el interesado respecto sus datos, a continuación se hablará sobre éstos:

- El artículo 12 explica que el interesado tiene el derecho de saber para qué motivo se están usando sus datos. Para esto, el responsable de tratamiento debe de tener los métodos apropiados de poder transmitirle esa información al interesado de forma clara, concisa y de fácil acceso.

- Acorde al artículo 13, cuando existan datos personales del interesado, el responsable de tratamiento debe proveer al interesado la siguiente información: la identidad y

formas de contactar al responsable de tratamiento y/o a su representante, los datos del delegado de protección de datos, los motivos por el cual se procesan los datos personales y su base legal, el periodo por el cual la información será guardada, la existencia del derecho a poder acceder, rectificar, restringir o borrar la información personal.

- En el artículo 14 se explica que, en caso de que el responsable de tratamiento obtenga información personal no haya sido dada por el interesado, el responsable del tratamiento deberá de entregar la información mencionada en el artículo 13.

- El artículo 15 dice que el interesado tiene el derecho de obtener confirmación del responsable de tratamiento acerca de que si existe procesamiento de su información personal, el interesado puede pedirle la información descrita en el artículo 13 al responsable de tratamiento.

- El artículo 16 explica que el interesado tiene derecho a que el responsable de tratamiento rectifique cualquier información errónea con respecto a él. De igual manera, el interesado tiene el derecho de tener información personal incompleta, completada

- El artículo 17 le da al interesado el derecho de exigir al responsable de tratamiento que se borre su información personal sin retrasos.

- El artículo 18 establece que el interesado tiene el derecho de restringir el procesamiento de su información si la veracidad de los datos del interesado es puesta en duda por el mismo interesado, o si el responsable de tratamiento ya no requiere la información personal para ser procesadas, pero sí por motivos de defensa o legales; de igual manera puede restringir su procesamiento si el interesado objeta, acorde al artículo 21.

- El artículo 19 estipula que en caso de que el interesado haya hecho una rectificación, una restricción o haya pedido que su información sea borrada, el responsable del tratamiento deberá de notificar a todos los destinatarios que hayan recibido dichos

datos personales, además de que deberá de notificar sobre dichos destinatarios al interesado, si él lo desea.

- En el artículo 20 se estipula que el interesado tiene el derecho de recibir sus datos personales en un formato estructurado y lectura mecánica, además de poder transmitir sus datos a otro responsable de manera fácil, mientras sea posible.

- El artículo 21 le otorga el derecho al interesado de poder objetar al procesamiento de su información y, en el caso de procesamiento por motivos de mercadotecnia, su información personal siempre dejará de ser procesada. En caso de motivos científicos, el procesamiento seguirá solamente si la información del interesado es vital para la investigación. El responsable de tratamiento tendrá derecho a prórroga si demuestra motivos legítimos para seguir con el procesamiento de los datos del interesado.

2.2.1.3. Responsable del Tratamiento y Encargado del Tratamiento

"Teniendo en cuenta la naturaleza, el ámbito, el contexto y los fines del tratamiento así como los riesgos de diversa probabilidad y gravedad para los derechos y libertades de las personas físicas, el responsable del tratamiento aplicará medidas técnicas y organizativas apropiadas a fin de garantizar y poder demostrar que el tratamiento es conforme con el presente Reglamento. Dichas medidas se revisarán y actualizarán cuando sea necesario"- (Art. 24 en [5]).

Dicho esto, el responsable del tratamiento debe de implementar medidas tanto tecnológicas como organizacionales para asegurar que la información personal sean procesadas para un motivo en específico. Estas medidas incluyen:

- La seudonimización y cifrado de los datos personales

- La capacidad de ofrecer confidencialidad, integridad, disponibilidad y resiliencia de los equipos y servicios

Dentro de las responsabilidades también cabe la notificación de violación de seguridad, en donde el responsable del tratamiento deberá de notificar a los interesados afectados solamente si el responsable del tratamiento no ha logrado mitigar la violación de ataque ni haber disminuido el riesgo de que se infrinjan los derechos del interesado.

2.2.1.4. Delegado de Protección de Datos

El responsable del tratamiento y el encargado del tratamiento designarán a un delegado protección cuando: el tratamiento sea llevado a cabo por un organismo o autoridad pública, el tratamiento de los datos requiera de una observación habitual y sistemática a gran escala de los interesados, el tratamiento consista en una gran cantidad de datos personales.Tanto el responsable del tratamiento como el encargado del tratamiento deben asegurarse que el delegado de protección de datos pueda hacer sus tareas de manera correcta y respaldar sus acciones, facilitando los recursos necesarios para el desempeño correcto, además de que se le debe de involucrar en todas las cuestiones referentes a la protección de datos personales.

El delegado debe de poseer amplio conocimiento sobre las leyes de protección de datos y sus formas de implementarlas. De tal forma, puede cumplir las siguientes funciones:

- Informar al responsable y encargado, además de los empleados que hagan el tratamiento de datos las obligaciones que tienen a partir de esta regulación.
- Asesorar cuando se requiera un análisis respecto a la protección de datos que se lleva a cabo.
- Resolver inquietudes de los interesados.
- Ser el puente entre la organización y las autoridades, además de cooperar con la autoridad de control.

Cabe destacar que las acciones del delegado de protección de datos siempre estarán acorde a la protección de datos de los interesados que tiene, y, no puede ser sancionado o penalizado ni tener conflicto de interés por cumplir sus funciones.

2.2.2. Privacidad y Analíticas de Aprendizaje

Hasta el momento, se ha hablado de la motivación y de las regulaciones que se tienen que cumplir para que un sistema que maneje, procesa o, como lo define la GDPR trate con la información personal de un tercero, sin embargo, no se han hablado de las pautas específicas que se deben de seguir para que una plataforma de análisis de aprendizaje sea segura y confiable respecto al público.

En [17] se explica que la aceptación a cualquier tecnología usando información en el ámbito de la educación depende de que el interesado esté consciente de las consecuencias que implica el usar la tecnología, y en específico, las plataformas de analíticas de aprendizaje, la validez y relevancia de los resultados obtenidos y saber la forma en la que sus datos son recolectados, procesados y compartidos.

El mayor problema que tiene el área de analíticas de aprendizaje radica en que, dado a su complejidad de recolectar datos y de sus procesos de análisis algorítmicos,no es trivial comunicar a los involucrados, como lo son los estudiantes, profesores, administradores o terceros como las autoridades educativas o los padres de los estudiantes, el cómo se recolectan los datos,qué datos se requieren para poder hacer el análisis y qué tan confiables son los resultados que entregó el análisis. Es por este motivo, que proponen abordar este tema desde un punto de vista institucional tomando en cuenta ámbitos éticos legales y de privacidad. De esta manera, se desarrolló DELICATE, una lista de ocho puntos que sirven de apoyo para los implementadores que les ayude a analizar riesgos de seguridad que pueden llegar a surgir debido a los procesos del manejo de datos y, cómo lidiar con ellos. A continuación se presentan los ocho puntos:

Determinación Por qué requiere aplicar las análitcas de aprendizaje, cuál es tu aportación y cuáles son los derechos de los interesados.

Explicar Se necesita ser explícito respecto los motivos y objetivos, además de decir por cuánto tiempo los datos serán guardados y quién tiene acceso a los datos.

Legitimar Porqué tienes el derecho de tener esos datos, qué datos ya posees y por qué no se son suficientes, y porqué tienes el derecho de guardar más datos

Implicar Involucra a todos los individuos que les concierne además de los interesados. Sé claro respecto a las preocupaciones de la privacidad de los interesados, además provee el acceso a los datos personales y educa al personal.

Consentimiento Establece un contrato con los interesados. Pide el consentimiento a los interesados de poder recolectar sus datos definiendo preguntas claras y entendibles, con respuestas de sí/no, además de ofrecerle al interesado la posibilidad de que dejen de recolectar su información sin repercusión alguna.

Anonimiza Procura que el interesado no sea identificado por sus datos, para esto, haz que los datos estén anonimizados lo más posible.

Técnico Instaura procedimientos para garantizar la privacidad del interesado: monitorea regularmente quién tiene acceso a los datos y, en caso de que las analíticas cambien, actualiza el consentimiento de los interesados. Finalmente, asegurate de que el lugar en dónde se almacenan los datos sigue los estándares internacionales de seguridad.

Externo En caso de que se colabore con proveedores externos, se tiene que cerciorar que también cumplan con las reglas organizacionales y nacionales. Para esto, el contrato debe de estipular claramente las responsabilidades respecto a la seguridad de datos y que los datos sólo deberán usarse para los servicios definidos y no para otros motivos.

2.3. Trabajos Relacionados

Una vez visto el problema que es difícil de resolver, dado a que se tiene que tener en consideración el factor humano, tanto de los usuarios como los responsables de la gestión de los datos, es necesario ver qué esfuerzos se han realizado por resolver este problema.

2.3.1. BBLAP

Los datos de aprendizaje reflejan las actividades hechas por los alumnos mientras aprenden. Con la gran cantidad de instituciones educativas y organizaciones, las diferentes implementaciones de plataformas de aprendizaje se vuelven inevitables. Por este motivo, se requiere tener un estándar para los datos y, a pesar de que existen estándares como el IMS Caliper y Tin Can Experience, los cuales han ayudado a reducir la carga de interoperabilidad en conjunción a los silos de datos de aprendizaje denominados Learning Record Store (LRS), sigue existiendo dificultad para la interoperabilidad sin limitantes. Es por eso que en [2] se propone la plataforma de analíticas de aprendizaje basada en Blockchain (BBLAP, por sus siglas en inglés). Sobre su plataforma, proponen smart contracts que contengan un los permisos al acceso a la información, su propietario y un mapeo entre los dos valores, de ésta manera, se cuentan con tres tipos de contratos:

Contratos de los Proveedores de aprendizaje Es el contrato encargado de controlar cómo las organizaciones e instituciones se vuelven proveedores autorizados en el Blockchain, deben de tener implementaciones para comunicarse y poder acceder a la información que las instituciones guardan.

Contratos de Estudiantes Este contrato representa la prueba de existencia de los datos de aprendizaje del estudiante en la plataforma del proveedor. Contiene la información del propietario, además de la dirección de la base de datos del proveedor, un hash de los datos de aprendizajes esperados y una lista de accesos para otros proveedores de aprendizaje.

Contrato Índice Este contrato contiene las relaciones entre los contratos de estudiante y los proveedores de aprendizaje, de esta manera, se mantiene un historial dentro de la Blockchain.

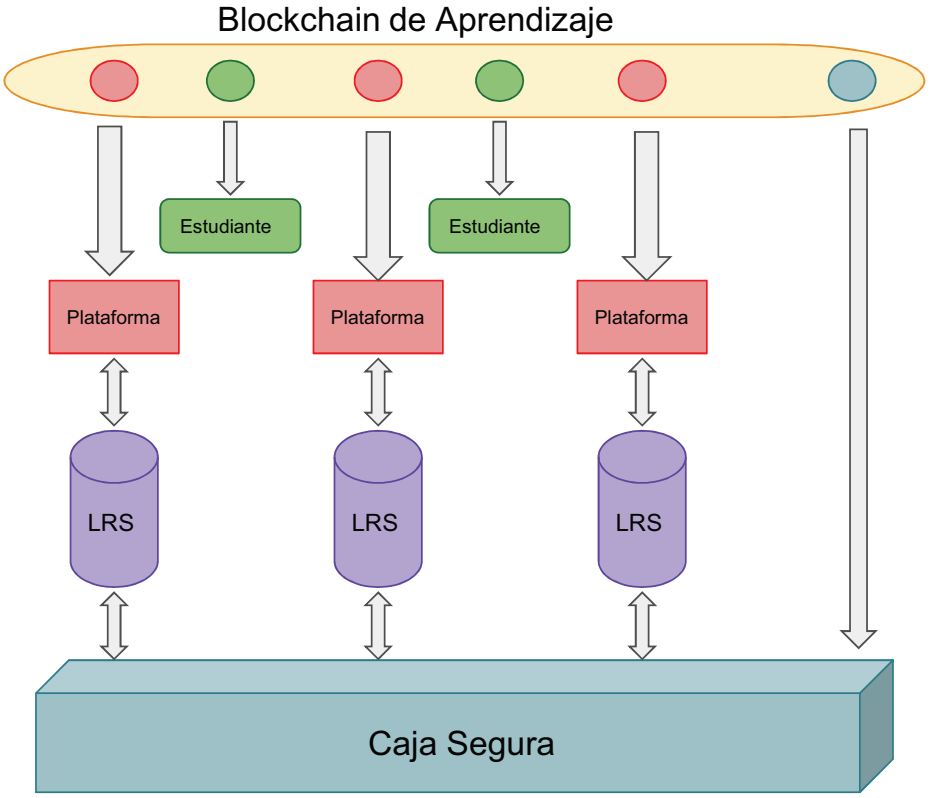

FIGURA 2.5: Diseño de la plataforma BBLAP. Fuente: [2] et al 2018

El diseño de esta plataforma gira en torno a la creación, adición y captación de los datos a través de Blockchain. El contenido de los bloques apunta a la información de aprendizaje, mientras que los nodos son los proveedores de la información y los estudiantes. Las actividades de aprendizaje quedarán registradas en el Blockchain.En este diseño, es requerido que los proveedores de la plataforma estén forzosamente como nodos en el Blockchain, y los estudiantes tienen la opción de no mantener un nodo, usando la plataforma como intermediario y que los proveedores sean los encargados de crear la cuenta del estudiante. De igual manera, se propone tener una API para la interacción a la Blockchain, para una comunicación más sencilla. También se propone una herramienta diseñada para mantener las interacciones seguras entre los LRS y los distintos proveedores. Para esto, se establecen las comunicaciones dentro de la caja segura que está ligada a todas las LRS. Dentro de

esta caja, se asegura de homogeneizar las bases de datos de los proveedores y, se mantiene una conexión a la Blockchain para verificar todas las ejecuciones de peticiones desde la caja segura, provienen de estudiantes o proveedores autorizados.

2.3.2. Hawk

Surgiendo ante la premisa de que el Blockchain se puede caracterizar como un intermediario, que puede ser confiado por su disponibilidad e integridad, pero no por su privacidad, Kosba *et al.* [18] presentan Hawk, un framework para crear y construir smart contracts con un énfasis la privacidad que no requiera de un especialista en criptografía para desarrollar un smart contract. Esto funciona gracias a que el compilador de Hawk se encarga de compilar el programa a un protocolo criptográfico. Para esto, el compilador divide el programa Hawk en tres piezas, de tal manera que el programa de Blockchain es ejecutado por todos los nodos; el programa que los usuarios ejecutarán y el programa el cual un tercero especial denominado el administrador podrá ejecutar. De igual manera, el programa resultante se divide en dos partes: La porción privada, la cual se encarga de guardar los datos y monedas. La otra parte, la porción pública, es la que no puede visualizar ni manipular los datos ni las monedas.

2.3.2.1. Seguridad

La seguridad de Hawk se basa en dos aspectos: El primero, establece la privacidad en la Blockchain, ya que, a menos de que los individuos involucrados en el contrato lo decidan, las partes privadas como los datos y las monedas no son mostradas al público debido a que dichas partes están ocultas por medio de criptografía. Si el primer aspecto protege a los individuos de todo aquel ajeno al contrato, el segundo aspecto protege a los individuos de sí mismos. Basándose en que ambos individuos actúan de manera egoísta para maximizar su beneficio financiero y que pueden desviarse del protocolo de manera arbitraria o inclusive abortar su parte del contrato, la seguridad contractual es una noción multifacética

de no solo confidencialidad y autenticidad, sino también de equidad financiera frente a los intentos de trampa o aborto.

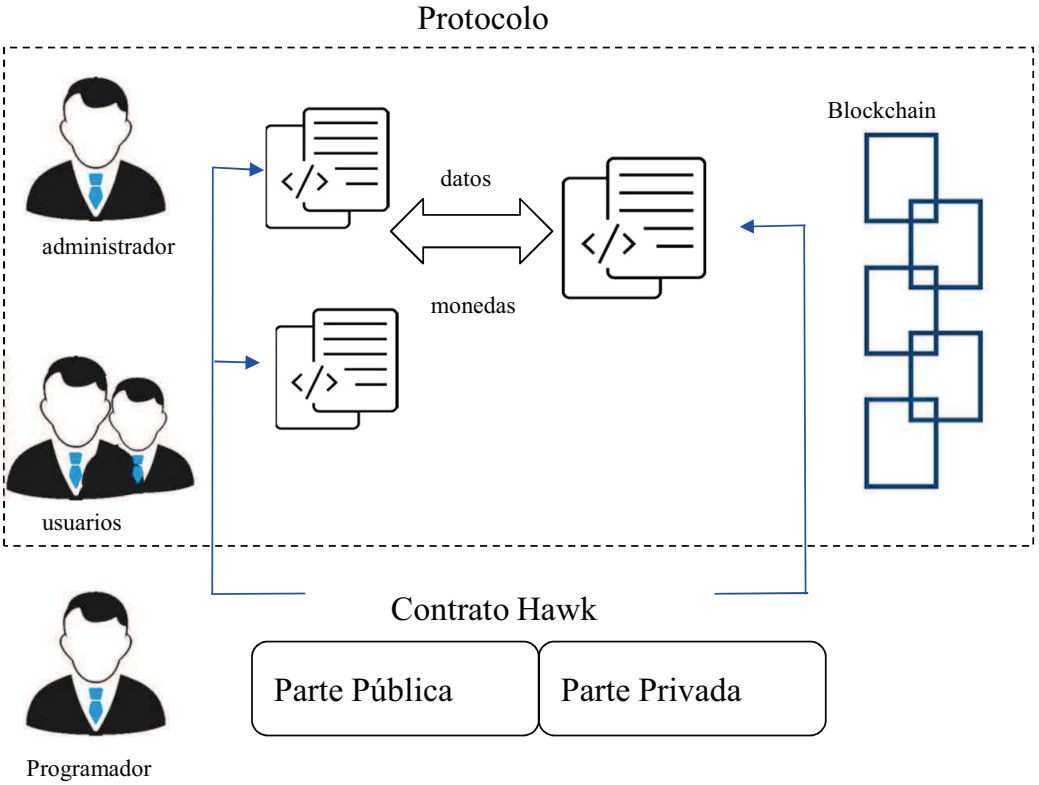

FIGURA 2.6: Contrato de Hawk. Fuente: Kosba et al 2016

De igual manera, un elemento importante en el diseño de Hawk, es el del administrador, un tercero especial que puede ver los valores de entrada de los usuarios y se le confía que no revele los datos de los usuarios. Sin embargo, al administrador no se le debe de tratar como un individuo de confianza, ya que inclusive el administrador puede desviarse del protocolo o coludirse con uno de los individuos del contrato, por lo tanto, el administrador no puede afectar la ejecución correcta del contrato y en el caso de que el administrador aborte el contrato, se le dará una penalización y los usuarios afectados obtendrán una compensación. Cabe mencionar, que al administrador debe de tener equipo de cómputo especializado para dar regiones privadas memoria desde el procesador, tal como el Intel SGX.

2.3.3. Desventajas de los trabajos relacionados

Si bien la plataforma de analíticas de aprendizaje basada en Blockchain propuesta por Ocheja *et al.* [2] da un gran aporte a la privacidad de los alumnos, ya que los incluyen en el diseño y les da una lista de permisos de qué otras plataformas pueden ver su información y que ellos mismos pueden editar, todavía tiene áreas de oportunidad respecto a la privacidad de los alumnos. El primer punto que se puede ver como desventaja es que los permisos de visualización a los datos de aprendizaje de un alumno, se aplican a toda la plataforma de análiticas de aprendizaje y no a un individuo, lo cual hace que dicha información pueda ser vista por cualquier usuario de la plataforma. El segundo punto es más una omisión, ya que la plataforma que maneja los datos de aprendizaje del alumno no tiene por qué pedirle autorización al alumno de acceder a sus datos negándole su derecho a la privacidad. El tercer punto recae en la carencia de temporalidad de los permisos de escritura, así que si un alumno no decide revocar un permiso de autorización a otra plataforma de análiticas de aprendizaje, esa plataforma tendrá acceso permanente a la información del alumno. Por otra parte, el enfoque de esta plataforma es proveer interoperabilidad entre distintas plataformas de análiticas de aprendizaje y no la preservación de la privacidad de sus alumnos.

Respecto al framework Hawk, su aporte principal radica en solucionar el problema de privacidad que tiene el Blockchain actualmente, sin embargo, su forma de resolverlo viene con ciertas desventajas. La primera desventaja recae en el rol del administrador, ese individuo tercero que puede ver los datos privados de los usuarios del Blockchain y, que bajo ningún motivo es completamente confiable y, si bien Hawk provee mecanismos de penalización hacia el administrador, no garantiza que el administrador siempre sea honesto. Por este motivo, Hawk no posee ventaja alguna de guardar los datos en el Blockchain contra tener un servidor seguro que se apoye del Blockchain para el manejo de datos.

2.4. Student Progress Snapshot

El Student Progress Snapshot es un dashboard de analíticas de aprendizaje el cual permite a los docentes analizar la actividad académica de sus estudiantes en cursos de Moodle desarrollado en el lenguaje PHP. Este proyecto toma en consideración las preocupaciones de seguridad y privacidad en torno a los datos personales de los alumnos;es por este motivo que en su diseño se busca incorporar mecanismos de seguridad tecnológicos que permitan la protección y restricción a la visualización de los datos, acorde a los puntos de Implicar, Consentimiento y Técnico de la lista DELICATE. De esta forma, la plataforma preservará los derechos de los estudiantes estipulados en el Reglamento General de la Protección de Datos.

Este proyecto aborda la implementación la seguridad de los datos con dos enfoques: la seguridad e integridad de los datos y los derechos de los estudiantes en respecto a sus datos. De igual manera, existen dos mecanismos de seguridad que trabajan en conjunción para brindarle al sistema la seguridad y la simpleza de la gestión de permisos: el mecanismo de encriptación a los datos y la implementación de Blockchain para la gestión de los permisos a acceso a los datos.

Con el diseño mostrado en la figura, se tiene planteado que la protección de los datos se haga mediante la encriptación de la base datos, mientras que la gestión de permisos será delegada al Blockchain. Esto se hace ya que en ningún caso, el Blockchain no puede ser usado para la proteger la privacidad de los estudiantes [19], sin embargo, dado a la inmutabilidad y la permanencia de los datos dentro del Blockchain, se vuelve la tecnología idónea para el almacenamiento de los permisos a los datos.

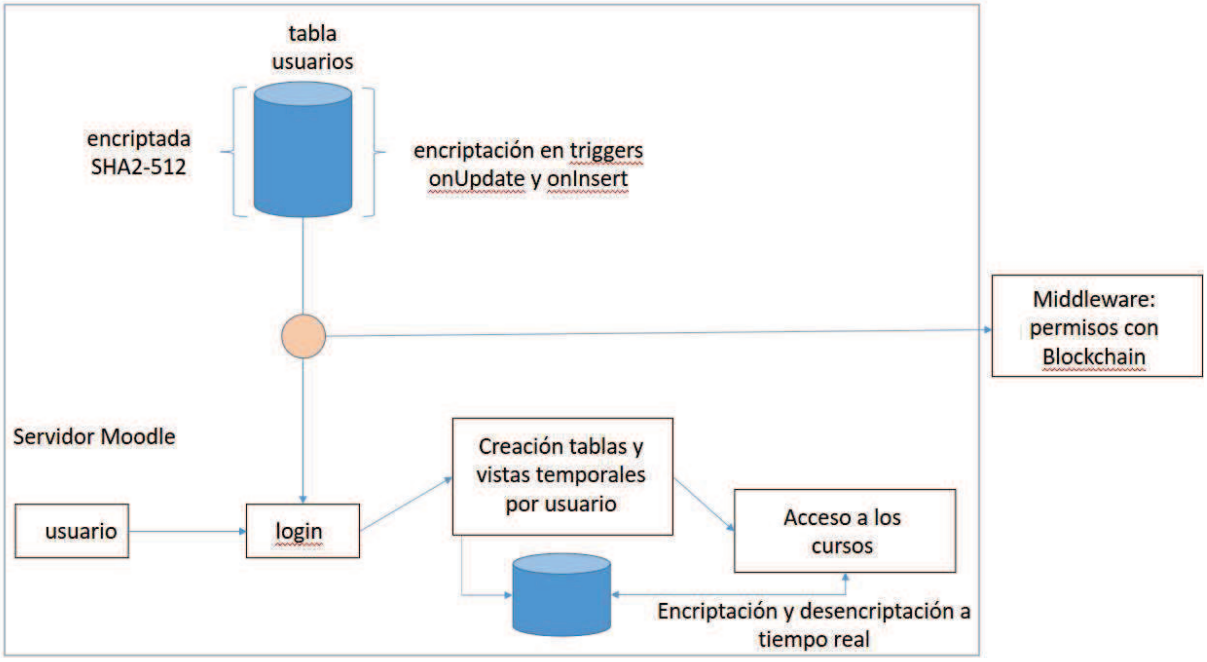

FIGURA 2.7: Esquema de la plataforma Student Progress Snapshot.

Capítulo 3

Trabajo Realizado

El objetivo de este trabajo, es implementar el módulo de Blockchain que está propuesto el proyecto de "Student Progress Snapshot" de tal manera que los alumnos puedan proveer permisos de acceso a sus datos. A continuación, se explicarán el cómo está construida la infraestructura, seguido de cómo está estructurado el smart contract usado. Finalmente, se explicarán las tecnologías necesaria y sus funciones para su implementación.

3.1. Infraestructura

Para empezar, por motivos de funcionalidad se escogió Ethereum como la red de Blockchain, dada su versatilidad al poder de implementar programas que no estén forzosamente relacionados a criptomonedas mediante el uso de smart contracts y por tener la capacidad de que los smart contracts pueden comunicarse con otros contratos.

Dado a que la plataforma Student Progress Snapshot se encuentra desarrollada en el lenguaje PHP y la manera más óptima para comunicarse con la Blockchain de Ethereum es W3.js y sus diversas implementaciones , se decidió crear un middleware que funcione como intermediario entre el Blockchain de Ethereum y el Student Progress Snapshot.

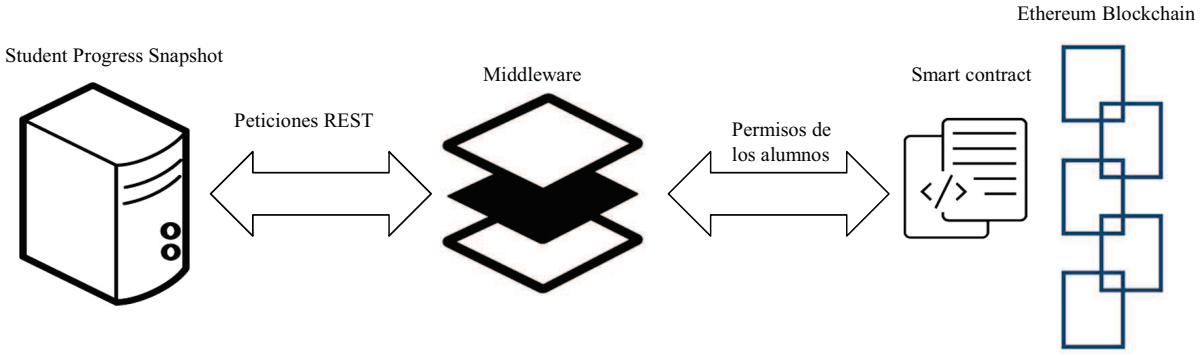

FIGURA 3.1: Esquema de la interacción entre el Blockchain y el Student Progress Snapshot.

Por cuestiones de diseño, seguridad y con el motivo de que no se desea hacer un pago monetario cada vez que se requiera añadir o editar un permiso de un alumno, se decidió tener una instancia privada de una Blockchain de Ethereum, de esta manera, no se requiere ningún pago para poder hacer el procesamiento y limita la cantidad de nodos que pueden llegar a ver los permisos establecidos por los alumnos.

3.2. Permisos usando Smart Contracts

Sin duda el elemento más importante de este trabajo es el smart contract, ya que en este yace la capacidad de guardar y administrar los permisos de accesos dados por los estudiantes. Si bien el Blockchain mantiene sobre su implementación el smart contract usado, además de la bitácora de transacciones, no se podría resolver el problema de proveer una plataforma de analíticas de aprendizaje en la cual tanto los tutores legales, administradores, docentes y sobretodo alumnos, tengan la confianza de que los permisos que establecieron los estudiantes sean editados por los responsables de la plataforma.

Para el diseño del smart contract, se tomó en consideración dos aspectos: los derechos del alumno respecto a sus datos y las obligaciones que se tiene como responsable y procesador de los datos. Los derechos relacionados al desarrollo de este trabajo incluyen: el alumno tiene derecho a pedir que se borre su información, saber quién puede ver su información

personal, a recibir sus datos, a negarse dar información personal y modificarla. Por otra parte, para el desarrollo del smart contract, se tomaron como enfoque las siguientes responsabilidades que se tienen al manejar los datos personales: Se debe de proveer de una fecha clara para el manejo y acceso a los datos del alumno, además de que debe de existir un consentimiento claro de que el alumno accedió a que sus datos personales sean analizados y consultar al alumno sobre quiénes pueden tener acceso. Además, se tuvo como objetivo secundario mantener la estructura sencilla, para que al momento de explicar el funcionamiento a los involucrados, haya menos riesgo de un malentendido.

3.3. Estructura

Para una gestión clara y sencilla de los permisos, se dividieron en dos diccionarios. Un diccionario contiene los permisos generales del alumno respecto al sistema y el otro diccionario tiene los permisos individuales que el alumno proporciona a los docentes. El diccionario de permisos generales del alumno existe para asegurar sus derechos y para demostrar de forma clara que existen los consentimientos del alumno. Este diccionario tiene como llave el identificador del alumno, tiene como contenido:

- La confirmación de que el alumno ha acepto los términos de privacidad del Student Progress Snapshot y la duración de estos términos. Esto en código se representa por un valor booleano y dos fechas en notación de Unix Epoch Time.
- La decisión del alumno respecto a que sus compañeros puedan acceder a su foto de perfil, esto representado por un valor booleano.
- El consentimiento del alumno respecto a que se le haga un seguimiento a las bitácoras que crea al usar la plataforma, de igual manera, se representa este valor por un booleano.

- El deseo del alumno respecto a que su información sea eliminada después de que haya concluido su curso académico, representado por un valor booleano.
- Saber si el alumno desea descargar su información personal, representado por un valor booleano.
- La petición del alumno explicando que desea que su información sea anonimizada, representada por un valor booleano.

El diccionario de permisos individuales sirve para llevar un control de qué docentes tienen acceso a qué información de qué alumno. Éste, tiene como llave el identificador del docente, mientras que su contenido es otro diccionario, que tiene como llave el identificador del alumno y como contenido tiene:

- El permiso del alumno de que el profesor pueda revisar sus datos.
- El tiempo de inicio y fin en el cual el profesor puede ver los datos del alumno.

Cabe mencionar, que cuando se agregue un alumno a los datos del instructor, éste puede observar los datos necesarios para que pueda ejercer su rol de profesor, como el nombre y apellido del alumno, a pesar de que el alumno haya denegado el acceso a sus datos personales. Esto se hace para que la plataforma de aprendizaje no se vea impedida por la implementación de seguridad y de esta manera, los datos personales, como la foto del alumno, su mail, entre otros, quedan salvaguardados.

A partir de estos diccionarios, se crea el algoritmo 3.1 para verificar si un docente puede acceder a los datos personales de un alumno.

El motivo de ésta verificación cuádruple es porque antes de que el docente pueda acceder a la información personal del alumno, se debe de verificar que el alumno haya estado de acuerdo a los términos de privacidad y que sigan vigentes esos términos, de lo contrario, cualquier acceso a la información del alumno implicaría una violación a sus derechos.

Algorithm 3.1 Algoritmo de verificación de accesos de los datos de un alumno

```
1: function VERIFICACIÓN(D, G, A, P, FIA, FTA, TA, FID, FTD, DA)
2:                          ▷ D: diccionario de docentes, G: diccionario general, A: alumno,
   P: docente, FIA: fecha inicial que el usuario permite que se acceda a su información
   en general, FTA: fecha de término en la cual se puede acceder los datos del usuario,
   TA: tiempo actual, FID: fecha de inicio la cual el docente puede acceder a los datos
   del alumno, FTD: fecha de término la cual el docente puede acceder a los datos del
   alumno, DA: permiso de acceso dado por el alumno al profesor.
3:     if (G{A:CP} es verdadero) then
4:         if (G{A:FIA} ≤ TA ≤ G{A:FTA}) then
5:             if (D{P:{A:FID}} ≤ TA ≤ D{P:{A:FTD}} then
6:                 if (D{P:{A:DA}} es verdadero) then
7:                     return Verdadero
8:                 end if
9:             end if
10:        end if
11:    end if
12:    return Falso
13: end function
```

Posteriormente, tiene que cerciorarse que el docente tenga el permiso del alumno y que esté dentro de la fecha establecida por el mismo alumno. Si no se cumpliearan esas condiciones, el docente no tiene permiso para acceder a los datos personales del alumno.

3.4. Implementación

Si bien el smart contract y su codificación son el elemento vital de este trabajo, éste sólo por sí mismo no resuelve el problema. Es por eso, que se tuvo que desarrollar la infraestructura antes mencionada. Dicho esto, a continuación se dará una descripción de cómo está implementado el Blockchain de Ethereum y el Middleware. Como ambiente de desarrollo se usó el sistema Linux Ubuntu 17.07

3.4.1. Implementación de Etherum

Para montar la red de Blockchain Ethereum, se utilizó “Go Ethereum” (Geth), una de las implementaciones originales del protocolo de Ethereum. Una vez que se haya instalado la

librería para ejecutar una red Ethereum. el primer paso es construir es construir el primer bloque del Blockchain llamado bloque Génesis, éste siempre codificado manualmente y generalmente no hace referencia a un bloque anterior, este archivo es de tipo JSON [20].

```
{
  "config": {
        "chainId": 0,
        "homesteadBlock": 0,
        "eip155Block": 0,
        "eip158Block": 0
    },
  "alloc"      : {},
  "coinbase"   : "0x0000000000000000000000000000000000000000",
  "difficulty" : "0x20000",
  "extraData"  : "",
  "gasLimit"   : "0x2fefd8",
  "nonce"      : "0x0000000000000042",
  "mixhash"    : "0x0000000000000000000000000000000000000000000000000000000000000000",
  "parentHash" : "0x0000000000000000000000000000000000000000000000000000000000000000",
  "timestamp"  : "0x00"
}
```

FIGURA 3.2: Ejemplo de contenido de un bloque Génesis. Fuente: [3]

Cabe remarcar que el primer bloque dicta las normas por las que se regirán los demás bloques ya que los parámetros que guarda incluyen el límite de gas que un bloque puede calcular antes de que se cree otro, la dificultad, que indica que tan computacionalmente complejo es minar una transacción, las cuentas de ethereum anteriores, entre otros. Para iniciar la red de Ethereum, solo es necesario ejecutar el comando: "geth –datadir ./datadir init Genesis.json" de esa manera, se iniciará el primer bloque. La forma en la que la red privada Ethereum funciona en este proyecto, requiere de otros parámetros al ser ejecutada, de esta manera, se pueda interactuar como con la misma.

```
Blockchain> geth --port 3000 --networkid 58343 --nodiscover --datadir=./bloc
 --maxpeers=0  --rpc --rpcport 8545 --rpcaddr 127.0.0.1 --rpccorsdomain "*"
--rpcapi "eth,net,web3,personal,miner"
```

FIGURA 3.3: Comando para ejecución de la red Blockchain.

A continuación, se explicarán los parámetros [21]:

Port Define el puerto por el que se puede acceder a la red Blockchain.

NetworkID Identificador de la red Blockchain. Este valor no puede ser de los ya utilizados para redes Ethereum públicas, dado a que la red Etherum usada es privada, cualquier valor numérico que no esté reservado puede ser usado.

Nodiscover Deshabilita el descubrimiento de nodos compañeros.

Datadir Directorio de datos en los que se ubican las cuentas externas y base de datos. maxpeers: Número máximo de nodos que se pueden conectar. En caso de que no haya conexión al exterior, ese valor debe ser 0.

rcp Habilita el servidor HTTP-RPC.

rcpport Define el puerto por el cual se interactua con el servidor HTTP-RPC.

rpcaddr Dirección del servidor HTTP-RPC.

rpcorsdomain Lista de dominios que puede aceptar peticiones.

rpcapi Las apis ofrecidas por servidor HTTP-RPC. La API de eth provee los métodos que permiten ver el estado de la Blockchain, las transacciones y las cuentas. La API de net abarca los métodos relacionados a los nodos de la red. La API de web3 da una interfaz para comunicarse con el RPC. El API de personal da métodos para la gestión de cuentas mientras que la API de miner da una forma rápida de hacer la minería de las transacciones [22].

3.4.1.1. Cuentas de Ethereum

Un aspecto que se debe de aclarar es que, a pesar de que ya se tenga una red de Ethereum ejecutándose, no hará nada a menos de que se añadan cuentas, tanto externas como contractuales. Para la solución de este trabajo, se tomó la decisión de que tanto los docentes como los alumnos no tienen la necesidad de tener una cuenta en el Blockchain Ethereum privado. Es por esto, que por diseño, solo se deberá de tener un contrato que gestione los permisos por cada instancia de análiticas de aprendizaje que llegue a estar

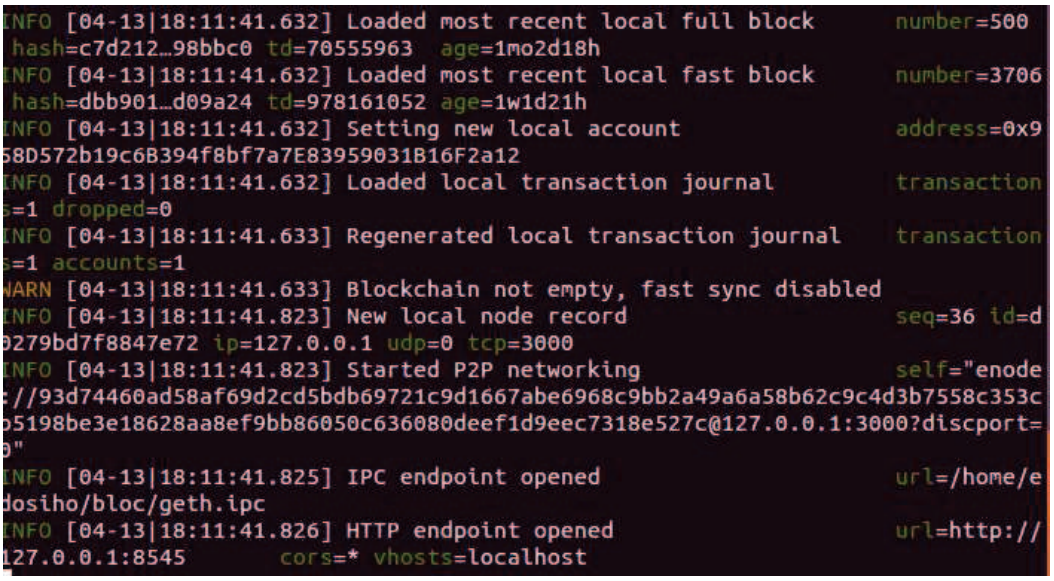

FIGURA 3.4: Nodo de Ethereum ejecutándose.

conectada al Blockchain. De esta manera, solo se tiene que acceder a un contrato para verificar los permisos. Además, al solo tener la cuenta institucional en el Blockchain, esa cuenta es la encargada de minar sus propias transacciones, trivializando los protocolos de cambio monetario que tiene de manera inherente el Blockchain.

3.4.1.2. Smart Contract Desarrollado

Si bien se habló sobre el diseño y el algoritmo del smart contract, hace falta señalar la forma en la que el contrato se agrega al Blockchain y los demás métodos que tiene implementados.

Hasta este momento, se ha omitido el hecho de un smart contract en Ethereum tiene múltiples lenguajes de programación en los que se puede escribir, esto se debe a que el lenguaje más usado para la creación de smarts contracts para Ethereum es Solidity, un lenguaje orientado a objetos de alto nivel inspirado en Python,C++ y Javascript, que su objetivo es la creación de smart contracts [23]. A pesar de que existan otros lenguajes como Serpent, un lenguaje de bajo nivel o Viper, que sigue en su fase experimental, la documentación oficial de Ethereum redirecciona a la documentación de Solidity y, que

el mismo creador de Serpent recomienda que se use Solidity para el desarrollo de smart contracts [24].

Es imperativo que para que un smart contract esté en la Blockchain, una cuenta externa debe de añadirlo. En este caso, la cuenta del Student Progress Snapshot es la que se encarga de subir ese contrato y en el momento en el que se encuentra en la Blockchain, la misma Blockchain le regresa a la cuenta externa la dirección del contrato, la cual sirve para que el resto de las cuentas se puedan comunicar con el contrato.

Un elemento que se crea al compilar el smart contract para que se añada al Blockchain, es la Interfaz Binaria de la Aplicación (Application Binary Interface, en inglés) del contrato. Esta interfaz es el estándar para interactuar con los contratos, tanto de contrato a contrato, como de cuentas externas al contrato, en una Blockchain de Ethereum [25]. Es menester que el ABI se preserve, de lo contrario, no se podrá interactuar de ninguna manera con el contrato. Esto se debe a que la codificación de los datos no está contenida en sí misma y por lo tanto, se necesita de un esquema para poder decodificar los datos. A continuación, se mostrarán los métodos que existen dentro del smart contract desarrollado:

AddLearner(learner,privacy, startDate, endDate,photo,logs,download, eraseData) Este método añade los permisos generales de un alumno al diccionario General.

AddPermissionToTeacher(teacher, learner,startDate,endDate, photo,data) el método que añade los permisos específicos de un estudiante a un docente.

CheckPermission(learner, teacher) Método que verifica si un docente tiene permiso de un alumno a ver sus datos personales.

GetPermissions(learner) Método que regresa los permisos generales de un alumno.

GetPhoto(learner, teacher) Método que verifica si el profesor puede ver la foto de perfil del alumno.

3.4.2. Middleware

La función de este middleware radica en permitir que el Student Progress Snapshot pueda interactuar de manera sencilla con el servidor de Ethereum, sin la necesidad de enviarle directamente las peticiones mediante RPC, ya que, como se mencionó antes, inclusive los desarrolladores de ethereum aconsejan en su documentación usar la librería de w3. Es por este motivo que se desarrolló un servidor capaz de comunicarse con la Blockchain Ethereum privada mediante una implementación basada en python de w3 llamada w3py y también comunicarse el Student Progress Snapshot mediante peticiones REST mediante un framework llamado Falcon, ya que este framework solo tiene los elementos básicos de un servidor web, haciendo énfasis para microservicios a larga escala con respuesta rápida, además de que está orientado a una arquitectura REST.

Todos los métodos que se llaman desde una petición REST, ya sea un PUT o un GET, se encargan solamente de recibir los parámetros necesarios para llamar a ejecutar el código de Web3py. Si bien existen la misma cantidad de métodos que hay en el smart contract, se cuenta con dos métodos extra:

compileContract() Compila el contrato para que pueda ser añadido al smart contract. Cabe señalar que el script del contrato reside en el servidor.

deployContract() Hace la transacción por parte de la cuenta externa poseída por la plataforma y guarda tanto el ABI como la dirección del contrato añadido en un archivo JSON.

Con la implementación de este middleware capaz de añadir y consultar los permisos de acceso a los datos personales de los estudiantes ubicados en el smart contract antes descrito, mediante consultas REST, se cumplieron los objetivos propuestos. Ya que al tener los permisos de acceso sobre la red Blockchain Ethereum, cada transacción queda guardada de manera permanente en el Blockchain, de tal forma que se mantiene un historial inalterable sobre los cambios hechos a los permisos dentro del smart contract;

además, gracias al diseño del smart contract, se es posible mostrar los permisos dados por el alumno de manera clara, con tan solo hacer una consulta al Blockchain. Estos objetivos en turno demuestran que la tecnología Blockchain es una solución viable para dar confianza a la plataforma, ya que provee a ésta los mécanismos de autorización y contabilidad.

Capítulo 4

Conclusiones

4.1. Conclusiones

La información que una persona llega a generar es importante y se le puede explotar para distintas aplicaciones, siendo las analíticas de aprendizaje una de éstas. Sin embargo, se tienen que respetar los derechos que los usuarios respecto a su información, tanto para que los mismos usuarios tengan confianza para que proveean sus datos como para no violar reglamentos legales, como es la GDPR.

A pesar de que esta regulación impone las normas a seguir, no existe método único para generar un sistema completamente confiable y seguro, ya que siempre existirá un factor de riesgo, tanto externo como interno; estos bien pueden ser ataques informáticos con el fin de acceder sin autorización a los datos o su eliminación o corrupción. Es por esto que se fija la atención a tecnologías emergentes para analizar su funcionamiento y determinar si pueden llegar a resolver las problemáticas que se tienen. Esto siendo el caso de este trabajo, el cual se propuso determinar si una implementación de Blockchain ayudaría a resolver el problema de la seguridad de los datos personales de los alumnos en una plataforma de analíticas de aprendizaje.

Tal como demostró la investigación y todo el trabajo realizado, una Blockchain actualmente no puede resolver por sí misma el problema de la seguridad de datos, ya que por diseño cualquier nodo que esté conectada a ella puede ver la información que lleguen a subir sus usuarios, lo cual niega completamente el propósito de seguridad y privacidad. Sin embargo, esto no implica que no sea útil, ya que gracias que dentro la Blockchain las transacciones hechas son inmutables, se puede tener una bitácora de confianza en la cual solamente se guarden los permisos de acceso, de esta manera, se torna complicado la manipulación al acceso y, se tiene una forma de presentar de manera clara los permisos de acceso a las autoridades y, si se usa en conjunto con mecanismos para la protección de datos como la encriptación, el sistema resultante será tanto seguro como confiable.

4.2. Trabajo a futuro

Si bien se cumplió el objetivo de poder tener un mecanismo que salvaguarda los permisos mediante el uso de Blockchain y smart contracts, sigue persistiendo el problema de privacidad que persiste en tecnología Blockchain, es por esto que se podría investigar y analizar la viabilidad de tener una Blockchain con los beneficios que ésta provee actualmente más la certeza de que los datos puedan permanecer seguros dentro de esta, a diferencia de lo que ocurre actualmente, que los datos deben de permanecer seguros en un servidor ajeno. Otro aspecto en el cual se puede trabajar es en la interoperabilidad entre plataformas de analíticas de aprendizaje y, que la Blockchain implementada por parte del Student Progress Snapshot sea usada en conjunción con otras instituciones para que éstas puedan compartir datos pertinentes usados para análisis, siempre y cuando los derechos de los alumnos se mantengan y que los permisos estipulados en este trabajo se apliquen igual a los externos.

Bibliografía

[1] Satoshi Nakamoto et al. Bitcoin: A peer-to-peer electronic cash system. 2008.

[2] Patrick Ocheja, Brendan Flanagan, and Hiroaki Ogata. Connecting decentralized learning records: a blockchain based learning analytics platform. In *Proceedings of the 8th international conference on learning analytics and knowledge*, pages 265–269. ACM, 2018.

[3] 0mkara. Ethereum private network configuration guide. https://gist.github.com/0mkara/b953cc2585b18ee098cd, 2017.

[4] Charles P Pfleeger and Shari Lawrence Pfleeger. *Security in computing*. Prentice Hall Professional Technical Reference, 2002.

[5] Reglamento general de protección de datos, 2016.

[6] L Johnson, R Smith, H Willis, A Levine, and K Haywood. The 2011 horizon report, 2011.

[7] Benjamin Herold. inbloom to shut down amid growing data-privacy concerns. *Education Week*, 21, 2014.

[8] Natasha Singer. Inbloom student data repository to close. *New York Times*, 21:2014, 2014.

[9] Aviv Zohar. Bitcoin: under the hood. *Communications of the ACM*, 58(9):104–113, 2015.

[10] Vitalik Buterin et al. Ethereum white paper. *GitHub repository*, pages 22–23, 2013.

[11] Remi Arpaci-Dusseau and Andrea Arpaci-Dusseau. Introduction to distributed systems, 2014.

[12] Accessing contracts and transactions. http://ethdocs.org/en/latest/contracts-and-transactions/accessing-contracts-and-transactions.html#rpc, 2016.

[13] Kevin Delmolino, Mitchell Arnett, Ahmed Kosba, Andrew Miller, and Elaine Shi. Step by step towards creating a safe smart contract: Lessons and insights from a cryptocurrency lab. In *International Conference on Financial Cryptography and Data Security*, pages 79–94. Springer, 2016.

[14] Massimo Bartoletti and Livio Pompianu. An empirical analysis of smart contracts: platforms, applications, and design patterns (2017), 2017.

[15] Rebecca Ferguson and Simon Buckingham Shum. Learning analytics to identify exploratory dialogue within synchronous text chat. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, pages 99–103. ACM, 2011.

[16] Cristóbal Romero, Sebastián Ventura, and Enrique García. Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1):368–384, 2008.

[17] Hendrik Drachsler and Wolfgang Greller. Privacy and analytics–it's a delicate issue. a checklist to establish trusted learning analytics. 2016.

[18] Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen, and Charalampos Papamanthou. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In *2016 IEEE symposium on security and privacy (SP)*, pages 839–858. IEEE, 2016.

[19] Daniel Amo, David Fonseca, Marc Alier, Francisco José García-Peñalvo, and María José Casañ. Personal data broker instead of blockchain for students' data privacy assurance. In *World Conference on Information Systems and Technologies*, pages 371–380. Springer, 2019.

[20] Genesis block. https://en.bitcoin.it/wiki/Genesis_block, 2017.

[21] Command line options. https://github.com/ethereum/go-ethereum/wiki/Command-Line-Options, 2018.

[22] Management apis. https://github.com/ethereum/go-ethereum/wiki/Management-APIs, 2017.

[23] Solidity. https://solidity.readthedocs.io/en/develop/index.html, 2016.

[24] Serpent. https://github.com/ethereum/serpent, 2016.

[25] Contract abi specification. https://solidity.readthedocs.io/en/develop/abi-spec.html, 2016.